

# 零交さ波分析装置について

## A SYSTEM FOR ANALYZING ZERO-CROSSING WAVES OF JAPANESE VOWELS

博士課程工学研究科電気工学専攻1年

石 田 義 久  
ISHIDA YOSHIHISA

### 1. まえがき

音声は文字と同様、人間の意志、思考を伝達する上で重要な役割をもっている。そこで、音声に関する研究は言語学や心理学をはじめとし、医学、音響工学など多方面の分野において古くから行なわれてきた。とくに、近年の電子計算機の発達による、電気工学、通信工学上の発展は著しく、音声を一連のパターン認識の問題として捉える領域にまで進展させている。

ところで、音声に関する研究には2つの側面がある。一つは、音声合成に関する研究で、これは、録音編集方式やダイナミックアナログ方式などの開発により、実用化が主たる研究課題となっている。一方、人間の声を機械によって自動的に識別し、適当な符号に変換しようとする、いわゆる音声認識に関する研究は現在発展途上にあり、これまで各種の方法が提案され、実験が行なわれてきた。この音声認識の究極的な目的は、人間と機械との間の能率的な情報交換にあり、そのプロセスとして、人間の音声処理機構の解明が目的となっている。また、応用例としては、電話の自動ダイヤリング、音声による計算機の自動制御、さらには自動翻訳機などが挙げられる。

さて、音声認識に対する研究の方向としては、先に報告したように<sup>1)</sup>、音声波の周波数スペクトル分析、零交さ波による分析などがある。周波数スペクトル分析法は、技術的に比較的簡単な方法であり、たとえば、十〜数十個の帯域フィルタによって周波数分析すればよい。ところが、この方法は、Heisenberg の不確定性原理に基づく周波数-時間分解能（いま、時間分解能を $\Delta t$ 、周波数分解能を $\Delta \omega$ とすれば、 $\Delta t \cdot \Delta \omega \simeq 1$  なる関係が存在する）によって、音声波の時間的変化を正しく捉えようとする場合には、広帯域のフィルタを必要とし、また各調波成分の分離を目的とするときには、狭帯域のフィルタが要求される。そこで、フィルタ群によるスペクトル分析の場合、目的（たとえば、母音のホルマント構造の分析を目的とするか、あるいは過渡音やわたりの特性解析を目的とするかなど）に応じて帯域巾を設定する必要がある。しかし、これを対象とする音声の特性によって自動的に行なおうとすると、かなり難しい問題であり、通常、音声認識の前処理としてフ

フィルタ群を用いる場合には、基本周波数以上の比較的広帯域のフィルタが使用される。したがって、各成分周波数がほぼ調和構造をなしている母音、有声子音などに対する高調波成分の分離は比較的困難なものとなる。

一方、零交さ波による音韻識別は、近年積極的に研究されるようになってきた方法で、つぎのような特徴をもっている。

まず、音声が発理パターンに変換されるので、音圧レベルに関する情報量が大幅に減少し、取扱いが容易となる。

つぎに、音声をクリッピングすることによって、ホルマント成分が強調されるようになるから<sup>2)</sup>、音韻識別における特徴抽出が容易となる。

さらに、フィルタ分析のように多数のフィルタを必要としないから、過渡音やわたりなどに対して、その時間的情報を正しく検出し得るようになる。

しかし、このような特徴をもつ反面、零交さ波は、同一帯域内にエネルギーの大きな成分が2個以上あると混変調を生じるので、必ずしも零交さ間隔がホルマントに対応しないという欠点をもっている。この混変調による影響を除去する方法として、帯域を2～3に分割し、各帯域ごとに零交さ波に変換するという方法が考えられている<sup>3)</sup>。

以上のように、周波数スペクトル分析法と零交さ波分析法には、それぞれ一長一短があり、音声認識の前処理としてどの方法が最も優れているかという問題は、今後の研究結果を待たねば解決しえない課題である。むしろ、音声の複雑多岐な性質に起因して、ある一つの分析法でもって音声のもつ全ての特徴を明らかにすることよりも、簡単ないくつかの手法を併用して、これらの総合的結果から判断を下すことのほうが、はるかに合理的であると考えられる。

このような目的で、筆者は、従来のフィルタ分析法<sup>1)</sup>に加えて、零交さ波分析装置を開発した。

本文で述べる零交さ波分析装置は、音声を3チャンネルの帯域フィルタによって音声帯域を分割してから、零交さ波に変換している。零交さ波から特徴抽出する方法としては、先に報告した<sup>4)</sup>積分器と自乗器から成る周波数測定回路を用い、零交さ間隔を実時間で周波数領域に変換している。

母音を識別するためのホルマント領域の分離回路は、演算器と論理回路とによって構成されており、第1ホルマント領域に対して5分割、第2ホルマント領域に対して6分割としている。分離回路の出力には、第1および第2ホルマント領域のそれぞれに対し、各サンプル時点においてただ一つのチャンネルのみ出力“1”が現われるようになっている。したがって、音声は、ホルマント周波数に対応したチャンネルのみ出力“1”をもつデジタルパターンに変換されるので、以後の処理が発理回路によって構成可能となる。本文は、これらについての詳細を述べたものである。

## 2. 零交さ波分析装置の構成

### 2-1. 概 要

図1は本装置のブロック図である。音声は、まず線形増巾器によって増巾され、AGC回路に入る。AGC回路の出力は、アナログ演算要素によって構成された3チャンネルの帯域フィルタ部

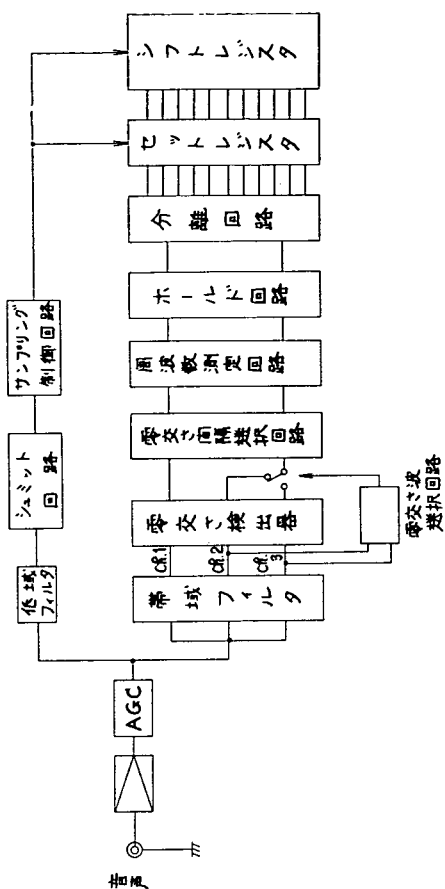


図 1

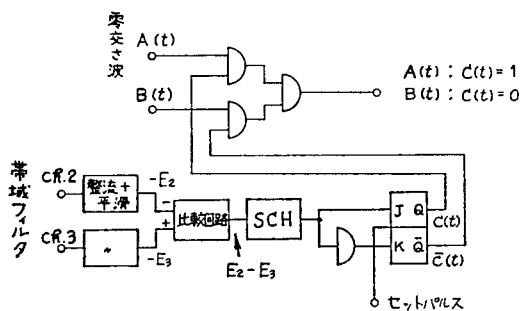


図 2

$E_2 \geq E_3$  ならば “1”

$E_2 < E_3$  ならば “0”

となり、フィルタ出力の大小が決定される。シュミット回路の状態は、各サンプル毎にフリップフロップに一時記憶され、より大きな出力をもったフィルタによる零交差波が選択されるように、論

[ch. 1)の帯域巾は320～900Hz、ch. 2) は950～1700 Hz、ch. 3) は1400～2700Hz] に入り、【音声帯域が分割される。第1ホルマント領域に対しては ch. 1 のフィルタが対応し、第2ホルマント領域に対しては ch. 2 または ch. 3 の何れかのフィルタが対応している。ただし、第2ホルマント領域に対応する ch. 2 および ch. 3 の入力、高域フィルタによって高域が強調されており、第1ホルマントとの混変調による影響を軽減している。また、AGC回路の出力は、サンプリング間隔をピッチ同期とするために低域フィルタに入り、基本波成分が抽出される。

つぎに、各フィルタの出力は、飽和要素とシュミット回路から成る零交差検出器によって零交差波に変換され、“0”、“1”の時系列論理パターンとして音声表現されることになる。この場合、第2ホルマント領域に相当する ch. 2 と ch. 3 のフィルタによる零交差波の選択は、それぞれの帯域フィルタの出力の大小によって決定される。図2はこの零交差波を選択する回路で、つぎのような動作を行う。まず、帯域フィルタ ch. 2 と ch. 3 の出力は、整流、平滑回路により直流電圧に変換され、振幅比較回路に入る。シュミット回路 (SCH) は、比較回路の出力を整形して、ch. 2 と ch. 3 のフィルタ出力の大小を決定する。すなわち、ch. 2 と ch. 3 のフィルタ出力の最大値をそれぞれ  $E_2$ 、 $E_3$  とすれば、整流回路で符号が反転され、比較回路の出力は、

$$E_2 - E_3$$

となる。したがって、シュミット回路の出力は

理回路から成る切換回路に入力される。

さらに、時系列論理パターンに変換された音声は、零交さ間隔選択回路を通り、積分器と自乗器から成る実時間周波数測定回路<sup>9)</sup>によって、時間領域から周波数領域に変換される。周波数測定回路の出力は、ホールド回路によって直流電圧に変換され、ほぼホルマント周波数に対応した電圧が得られることになる。

分離回路は演算器と論理回路から成り、ホールド回路の出力を各周波数帯域毎に分離する働きを行う。分離回路のチャンネル数は、第1ホルマント領域に対し5、第2ホルマント領域に対し6となっている。分離回路の出力は、ピッチに同期したサンプリングパルスによってセットレジスタに一時記憶され、さらに、ホルマント周波数の時間的変化を調べるために、シフトレジスタに順次記憶されていく。

このように、音声はホルマント周波数に対応したチャンネルのみ出力“1”をもつデジタル信号に変換されるので、以後の処理は論理回路の組合わせによって構成可能となる。

## 2-2 帯域フィルタ

音声を零交さ波に変換する場合、同一帯域内にエネルギーの大きな成分が2個以上あると、混変調によって零交さ間隔が必ずしもホルマント周波数に対応しないことがある。そこで、本装置では帯域フィルタによって音声帯域を分割してから、各帯域毎に零交さ波に変換している。図3は第1ホルマントを検出するための帯域フィルタで、IC化

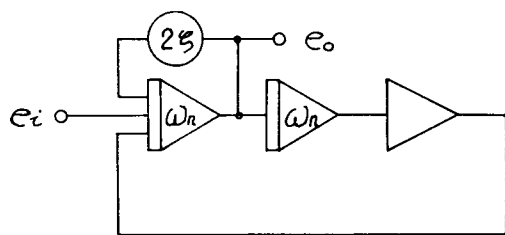


図 3

アナログ演算要素を用いて構成している。一方、第2ホルマントに対しては、5母音のホルマント周波数の分布範囲と第3ホルマントによる影響を考慮して、中心周波数の異なるフィルタを2チャンネル使用している。

図3に示すフィルタの入出力比  $e_o/e_i$  は

$$\frac{e_o}{e_i} = \frac{\omega_n s}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (1)$$

で表わされる。ここで、 $\zeta$  は制動比、 $\omega_n$  は固有角周波数である。また、インパルス応答の減衰特性  $h(t)$  は

$$h(t) = e^{-\zeta\omega_n t}$$

$Q = 1/2\zeta = \omega_n/2\pi B$  ( $B$ : 帯域巾) であるから

$$h(t) = e^{-\pi B t} \quad (2)$$

となる。したがって、フィルタの過渡応答は帯域巾に関係し、広帯域のフィルタは短時間で消滅するが、狭帯域のフィルタは過渡応答が長くなる。本方式のように、音声帯域の分割に使用されるフィルタは、一般にスペクトル分析に使われるフィルタに比べ帯域巾が広いので、音声の時間的変化を正しく検出し得るようになる。

o (a)

[illegible]

(3)

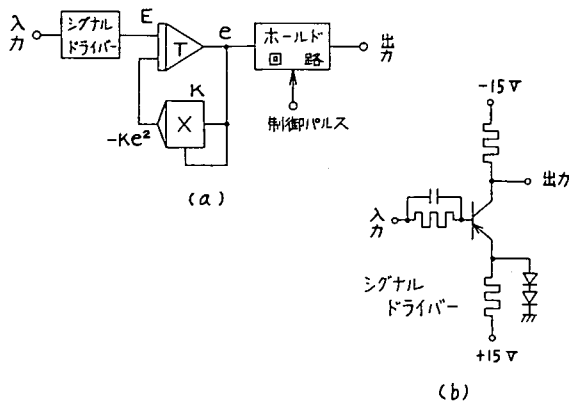


図 6

ただし、積分器の時定数を  $T$ 、初期値を  $E$ 、自乗器の比例定数を  $K$  とする。

とくに、 $\tau = T/KE$  に選べば

$$e = \frac{T}{K} \cdot \frac{1}{t} \propto \frac{1}{t}$$

となり、積分器の出力には、零交さ間隔  $t$  の逆数に比例した電圧が現われる。

ホールド回路は、単安定マルチバイブレータ  $MM_2$  によって制御されており、出力にはほぼホルマント周波数に対応した電圧が得られる。したがって、音声帯域を分割するフィルタに高速度掛算器を含むアナログ演算要素から成る同期化フィルタ<sup>1)</sup>を用い、上記のホールド回路の出力によって掛算器を制御すれば、つねに帯域フィルタの中心周波数をホルマント周波数の近傍に置くことができるようになる。このようにすれば、混変調による影響をさらに軽減できるものと思われるが、本方式については未だ検討を要するので、詳細はいずれ稿をあらためて述べたい。

#### 2-4 分離回路

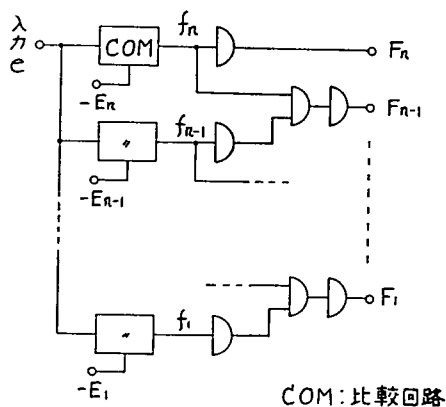


図 7

母音を識別するためのホルマント領域の分離回路は、図7に示すように演算器による比較回路と論理回路とによって構成されている。各演算器の出力  $f_i$  ( $i=1, 2, \dots, n$ ) は入力を  $e$  とすれば

$$e \geq E_i \text{ ならば } f_i = "0"$$

$$< E_i \text{ ならば } f_i = "1"$$

(ただし、 $0 < E_1 < E_2 < \dots < E_n$ である。)

となる。 $n$  の数は、第1ホルマント領域に対しては5、第2ホルマント領域に対しては6としている。

論理回路部の出力  $F_i$  ( $i=1, 2, \dots, n$ ) は

$$F_i = f_{i+1} \cdot \overline{f_i}$$

(ただし、 $f_{n+1} = 1$ )

となっている。ここで、

$$f_i = 0 \text{ ならば } f_{i-j+1} = 0$$

$$(j=1, 2, \dots, i)$$

$$f_{i+1} = 1 \text{ ならば } f_{i+k} = 1$$

$$(k=1, 2, \dots, n-i)$$

であるから、 $F_i = 1$  すなわち  $f_{i+1} = 1$  かつ  $f_i = 0$  なるとき

$$F_{i-j} = f_{i-j+1} \cdot \overline{f_{i-j}} = 0$$

$$F_{i+k} = f_{i+k+1} \cdot \overline{f_{i+k}} = 0$$

となる。よって、 $F_i = 1$  なるとき

$$F_{l(i \neq l)} = 0 \quad (i, l=1, 2, \dots, n)$$

となり、各サンプル時点において、ただ一つのチャンネルのみ出力“1”が現われる、分離回路の出力は、シフトレジスタに順次記憶され、ホルマント周波数の時間的变化を検出できるようになっている。

このように、音声はホルマント周波数に対応したチャンネルのみ出力をもつデジタル信号に符号化されるので、音韻識別装置は論理回路の組合わせによって構成可能となる。

### 3. 母音識別回路

本装置において、母音識別回路は図8に示すように論理回路によって構成されており、前章で述べた分離回路の出力状態から識別が行なわれる。同図において、定常性入力には、第1ホルマントを検出する周波数測定回路の出力と第2ホルマントを検出する周波数測定回路の出力の時間的变化(すなわち、ホルマント周波数の時間的变化に対応)が、いずれもある一定値以下にある場合“1”そうでない場合には“0”の信号が入る。母音部において、ホルマント周波数がほぼ一定値を保つ時間は、子音部分に比べ比較的長時間であるが、定常部(厳密には、音声がほぼ定常状態にあると考えられる区間)におけるホルマント周波数の変動巾は、人により、また発声する音韻等により種々

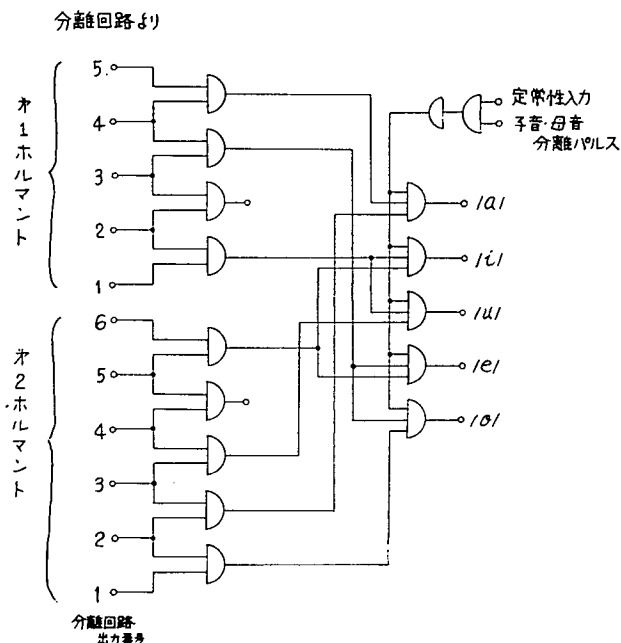


図 8

変化するので、上記定常性の決定（音声定常状態にあると見なすのに許容されるホルマント周波数の変動巾の決定）は実験結果と比較して行なっている。

子音、母音分離パルスは、子音区間と母音区間とを区別するための信号で、母音部であるときには“1”、子音部であるときには“0”となる。子音区間と母音区間とを分離する方法に、子音部と母音部における音圧レベルの差を利用する方式があるが、この方法は鼻音（/m、n、ŋ/）のように、レベル差の比較的小さい音韻に対しては誤動作を生ずることがある。そこで、本装置では、入力音声を低域フィルタと高域フィルタにパラレルに通し、その出力からピッチ成分の有無や、高域周波数成分を検出し、これらの組合わせから子音区間と母音区間の分離を行なっている。

#### 4. ピッチ抽出

音声の自然性（個人性、情緒性）を表わす基本的情報要素として重要な役割をもつと考えられるピッチ周波数の抽出は、原理的にホルマント周波数の抽出方法と同じであり、図9に示すような回

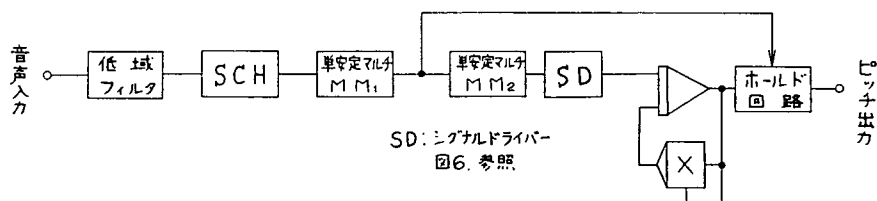


図 9



路構成になっている。

音声は低域フィルタを通して基本波成分が抽出され、シュミット回路により、パルスに波形整形される。つぎに、シュミット回路の出力は、ホールド回路を制御する単安定マルチバイブレータ  $MM_1$  に入り、時間軸推移用単安定マルチバイブレータに加えられる。

時間軸推移用単安定マルチバイブレータ  $MM_2$  の出力は、シグナルドライバ (SD) によって、0 と -15V に電圧変換され、積分器に入力される。ホールド回路は、 $MM_1$  によって制御され、積分器の出力を直流電圧に変換する役目を行う。

この結果、ホールド回路の出力には、ピッチ周波数に比例した電圧が、実時間で、しかも比較的高精度で得られることになる。筆者は、本装置を2章で述べた零交さ波分析装置と併用することによって、たとえば、対数増巾器などにより  $\log F_1/F_0$ 、 $\log F_2/F_0$  ( $F_0$ 、 $F_1$ 、 $F_2$  はそれぞれピッチ周波数、第1、第2ホルマント周波数を表わす) を求め、ホルマント周波数を正規化する方法について検討を進めている。

## 5. む す び

本稿は、積分器と自乗器から成る実時間周波数測定回路を用い、零交さ波からホルマント周波数とピッチ周波数を検出する方法について述べたものである。

音声認識において、決定的な認識機械を生み出すことは極めて困難である。そこで、簡単な幾通りかの判断機構の組合わせから、総合的判断を下すという方法も、一つの認識方法であると考えられる。このような目的に対して、本装置のように比較的簡単な回路方式による認識装置は有効であると思われる。

ここで、本装置の特徴を要約すれば、

- 周波数測定回路によって、零交さ波から直接特徴抽出を行なっている。
  - 母音を識別するためのホルマント領域の分離回路は、演算器と論理回路とによって構成されており、各サンプル点において、ホルマント周波数に対応したチャンネルのみ出力が現われるようになっている。
  - ピッチ周波数によって、ホルマント周波数を正規化する方式について検討している。
- などが挙げられる。

本稿では、認識系の構成を記述するにとどまった。現在、本装置は実験中であるので、認識率、得られたデータその他についての検討は、いずれ稿をあらためて述べたい。

おわりに、本研究にあたり終始御指導いただいた本学教授・後藤以紀博士、同教授・小川康男博士、同助教授・本多高博士に厚く御礼申し上げる。

## 参 考 文 献

- 1) 石田、小川他；“音声分析装置について”、明大工学部研究報告 No. 25, II-134
- 2) 平松；“音声の S. S. B.-clipping 波に含まれる情報” 音響学会誌 Vol. 18, No. 6. 1962 (P. 301~P. 309)

- 3) Robert W. A. Scarr : "Zero Crossings as a Means of Obtaining Spectral Information in Speech Analysis", IEEE Vol. AU-16, No. 2, 1968 (P. 247~P. 255)
- 4) 石田、小川 ; "ピッチ抽出装置と同期化フィルタについて", アナログ研究会資料 Vol. 11, No. 7, 1971 (P. 1~P. 11)